

## RAW SEQUENCE LISTING ERROR REPORT



COPY OF PAPERS  
ORIGINALLY FILED

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/947,081A

Source: OIPE

Date Processed by STIC: 08/09/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

**Raw Sequence Listing Error Summary**

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	<u>SERIAL NUMBER:</u> <u>09/847,081A</u>
<b>ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFT</b>		
1 <input type="checkbox"/> Wrapped Nucleic Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 <input type="checkbox"/> Misaligned Amino Numbering	The numbering under each 3 <sup>rd</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 <input type="checkbox"/> Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input checked="" type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 <input type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 <input type="checkbox"/> Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <input type="checkbox"/> Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.	



OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/847,081A

DATE: 08/09/2001

TIME: 12:44:02

Input Set : A:\Mo6314 sequence-TEXT.txt  
 Output Set: N:\CRF3\08092001\I847081A.raw

3 <110> APPLICANT: BAYER AG  
 5 <120> TITLE OF INVENTION: DNA encoding the tobacco phytoene synthase  
 7 <130> FILE REFERENCE: Le A 34 326  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/847,081A  
 C--> 10 <141> CURRENT FILING DATE: 2001-05-02  
 12 <160> NUMBER OF SEQ ID NOS: 10  
 14 <170> SOFTWARE: PatentIn Ver. 2.1  
 16 <210> SEQ ID NO: 1  
 17 <211> LENGTH: 1728  
 18 <212> TYPE: DNA  
 19 <213> ORGANISM: Nicotiana tabacum  
 21 <220> FEATURE:  
 22 <221> NAME/KEY: CDS  
 23 <222> LOCATION: (244)..(1566)  
 25 <400> SEQUENCE: 1  
 26 agaaaacccag aaagaacaac aggtttgct tcttggat gaggcattt gcctctgctt 60  
 28 gtgttaaggca aagtccggttc acttttttat atccgattttataatcggtt gaaatttagtg 120  
 30 gatagactct agtggatatacataaggattt gggtttttga taaaataggc tgagggtgaga 180  
 32 aggttaacata aaggaaagac aaaaacttgg gaattttttt agaccaccga ggtttcttgt 240  
 34 ttc atg agc atg tct gtc ttg ttg tgg gtt gtt tct ccc act tcc 288  
 35 Met Ser Met Ser Val Ala Leu Leu Trp Val Val Ser Pro Thr Ser  
 36 1 5 10 15  
 38 gag gtc tcg aat ggg aca gga ttg ttg gat tca gtc cga gaa gga aac 336  
 40 Glu Val Ser Asn Gly Thr Gly Leu Leu Asp Ser Val Arg Glu Gly Asn  
 41 20 25 30  
 43 cgc gtc ttt gta tca tcc agg ttc cta gct cga gat agg aat ttg atg 384  
 45 Arg Val Phe Val Ser Ser Arg Phe Leu Ala Arg Asp Arg Asn Leu Met  
 46 35 40 45  
 48 tgg aat ggg aga atc aag aaa ggt ggg aga caa agg tgg aat ttt ggc 432  
 50 Trp Asn Gly Arg Ile Lys Lys Gly Arg Gln Arg Trp Asn Phe Gly  
 51 50 55 60  
 53 tct tta att gct gat cca aga tat tca tgc ttg ggt gga tca aga act 480  
 56 Ser Leu Ile Ala Asp Pro Arg Tyr Ser Cys Leu Gly Gly Ser Arg Thr  
 57 65 70 75  
 59 gaa aag gga agc act ttc tct gta cag tcc agt ttg gtg gct agc cca 528  
 61 Glu Lys Gly Ser Thr Phe Ser Val Gln Ser Ser Leu Val Ala Ser Pro  
 62 80 85 90 95  
 64 gct gga gaa atg act gtg tca tca gag aaa aag gtg tat gat gtg gta 576  
 66 Ala Gly Glu Met Thr Val Ser Ser Glu Lys Lys Val Tyr Asp Val Val  
 67 100 105 110  
 69 tta aag cag gca gct tta gtg aag agg cag ctg aga tct acc gat gat 624  
 71 Leu Lys Gln Ala Ala Leu Val Lys Arg Gln Leu Arg Ser Thr Asp Asp  
 72 115 120 125  
 74 tta gaa gtg aag ccg gat att gtt gtt cca ggg aat ttg ggc ttg ttg 672  
 76 Leu Glu Val Lys Pro Asp Ile Val Val Pro Gly Asn Leu Gly Leu Leu  
 77 130 135 140  
 79 agt gaa gca tat gat cgt tgt ggc gaa gta tgt gca gag tat gca aag 720

Does Not Comply  
Corrected Diskette Needed

p. 4 Erroneous

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/847,081A

DATE: 08/09/2001  
TIME: 12:44:02

Input Set : A:\Mo6314 sequence-TEXT.txt  
Output Set: N:\CRF3\08092001\I847081A.raw

81 Ser Glu Ala Tyr Asp Arg Cys Gly Glu Val Cys Ala Glu Tyr Ala Lys  
82 145 150 155  
84 aca ttt tac tta gga acc aag cta atg acc cca gag aga aga aga gct 768  
86 Thr Phe Tyr Leu Gly Thr Lys Leu Met Thr Pro Glu Arg Arg Arg Ala  
87 160 165 170 175  
89 atc tgg gca ata tat gtg tgg tgc agg aga acg gat gag ctt gtt gat 816  
91 Ile Trp Ala Ile Tyr Val Trp Cys Arg Arg Thr Asp Glu Leu Val Asp  
92 180 185 190  
94 ggc cct aat gca tcc cac ata act ccg caa gct tta gat agg tgg gag 864  
96 Gly Pro Asn Ala Ser His Ile Thr Pro Gln Ala Leu Asp Arg Trp Glu  
97 195 200 205  
99 acc agg ctg gaa gat att ttc agt ggg cgg cca ttt gat atg ctt gat 912  
101 Thr Arg Leu Glu Asp Ile Phe Ser Gly Arg Pro Phe Asp Met Leu Asp  
102 210 215 220  
104 gct gct tta tcc gat act gtc tcc aga ttt cct gtt gat att cag cca 960  
106 Ala Ala Leu Ser Asp Thr Val Ser Arg Phe Pro Val Asp Ile Gln Pro  
107 225 230 235  
109 ttc aga gat atg att gaa gga atg cgt atg gac ttg tgg aaa tcc aga 1008  
111 Phe Arg Asp Met Ile Glu Gly Met Arg Met Asp Leu Trp Lys Ser Arg  
112 240 245 250 255  
114 tac aaa act ttc gat gag cta tat ctc tat tgt tac tat gtt gct ggt 1056  
116 Tyr Lys Thr Phe Asp Glu Leu Tyr Leu Tyr Cys Tyr Tyr Val Ala Gly  
117 260 265 270  
119 act gta gga ttg atg agt gtt cca gtt atg ggt att gca cct gaa tca 1104  
121 Thr Val Gly Leu Met Ser Val Pro Val Met Gly Ile Ala Pro Glu Ser  
122 275 280 285  
124 aag gca aca aca gag agt gta tat aat gct gct ttg gct tta ggg ctt 1152  
126 Lys Ala Thr Thr Glu Ser Val Tyr Asn Ala Ala Leu Ala Leu Gly Leu  
127 290 295 300  
129 gca aatcaa cta acc aat ata ctc aga gat gta gga gaa gat gcc aga 1200  
131 Ala Asn Gln Leu Thr Asn Ile Leu Arg Asp Val Gly Glu Asp Ala Arg  
132 305 310 315  
134 aga gga aga gta tac ttg cct caa gat gaa tta gca cag gca ggg ctc 1248  
136 Arg Gly Arg Val Tyr Leu Pro Gln Asp Glu Leu Ala Gln Ala Gly Leu  
137 320 325 330 335  
139 tcc gac gaa gac ata ttt gct gga aga gtg act gat aag ttg agg aac 1296  
141 Ser Asp Glu Asp Ile Phe Ala Gly Arg Val Thr Asp Lys Trp Arg Asn  
142 340 345 350  
144 ttt atg aag aaa caa att cag agg gcg agg aaa ttc ttt gat gag tca 1344  
146 Phe Met Lys Lys Gln Ile Gln Arg Ala Arg Lys Phe Phe Asp Glu Ser  
147 355 360 365  
149 gag aaa ggt gtc aca gaa ctg gac tct gct agt aga ttg cct ttg tta 1392  
151 Glu Lys Gly Val Thr Glu Leu Asp Ser Ala Ser Arg Trp Pro Val Leu  
152 370 375 380  
154 aca gcg ctg ctg ttg tat cgc aag ata ttg gac gag att gaa gcc aac 1440  
156 Thr Ala Leu Leu Tyr Arg Lys Ile Leu Asp Glu Ile Glu Ala Asn  
157 385 390 395  
159 gac tac aac aac ttc aca agg agg gct tat gtt agc aag cca aag aag 1488  
161 Asp Tyr Asn Asn Phe Thr Arg Arg Ala Tyr Val Ser Lys Pro Lys Lys

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/847,081A

DATE: 08/09/2001

TIME: 12:44:02

Input Set : A:\Mo6314 sequence-TEXT.txt  
 Output Set: N:\CRF3\08092001\I847081A.raw

```

162 400          405          410          415
164 ctt ctc acc ttg ccc att gct tat gca aaa tct ctt gtg ccc cct aat    1536
166 Leu Leu Thr Leu Pro Ile Ala Tyr Ala Lys Ser Leu Val Pro Pro Asn
167          420          425          430
169 aga act tcc tct cca cta gca aag aca tga atgaagtagt tgagtcaatg    1586
171 Arg Thr Ser Ser Pro Leu Ala Lys Thr
172          435          440
174 agtattatac actaaagaaa ctcaggtact tgtaaatgag atatctttg ctaaatgtgt 1646
176 atcatcaaaa gtagattgta aattcaatat gacaatctct tggtagaata ttttctccac 1706
178 actcatcaaaa ccctcaagtg ag                                1728
181 <210> SEQ ID NO: 2
182 <211> LENGTH: 440
183 <212> TYPE: PRT
184 <213> ORGANISM: Nicotiana tabacum
186 <400> SEQUENCE: 2
187 Met Ser Met Ser Val Ala Leu Leu Trp Val Val Ser Pro Thr Ser Glu
188   1           5           10          15
190 Val Ser Asn Gly Thr Gly Leu Leu Asp Ser Val Arg Glu Gly Asn Arg
191          20          25          30
193 Val Phe Val Ser Ser Arg Phe Leu Ala Arg Asp Arg Asn Leu Met Trp
194          35          40          45
196 Asn Gly Arg Ile Lys Lys Gly Gly Arg Gln Arg Trp Asn Phe Gly Ser
197          50          55          60
199 Leu Ile Ala Asp Pro Arg Tyr Ser Cys Leu Gly Gly Ser Arg Thr Glu
200   65          70          75          80
202 Lys Gly Ser Thr Phe Ser Val Gln Ser Ser Leu Val Ala Ser Pro Ala
203          85          90          95
205 Gly Glu Met Thr Val Ser Ser Glu Lys Lys Val Tyr Asp Val Val Leu
206          100         105         110
208 Lys Gln Ala Ala Leu Val Lys Arg Gln Leu Arg Ser Thr Asp Asp Leu
209          115         120         125
211 Glu Val Lys Pro Asp Ile Val Val Pro Gly Asn Leu Gly Leu Leu Ser
212          130         135         140
214 Glu Ala Tyr Asp Arg Cys Gly Glu Val Cys Ala Glu Tyr Ala Lys Thr
215   145         150         155         160
217 Phe Tyr Leu Gly Thr Lys Leu Met Thr Pro Glu Arg Arg Arg Ala Ile
218          165         170         175
220 Trp Ala Ile Tyr Val Trp Cys Arg Arg Thr Asp Glu Leu Val Asp Gly
221          180         185         190
223 Pro Asn Ala Ser His Ile Thr Pro Gln Ala Leu Asp Arg Trp Glu Thr
224          195         200         205
226 Arg Leu Glu Asp Ile Phe Ser Gly Arg Pro Phe Asp Met Leu Asp Ala
227          210         215         220
229 Ala Leu Ser Asp Thr Val Ser Arg Phe Pro Val Asp Ile Gln Pro Phe
230   225         230         235         240
232 Arg Asp Met Ile Glu Gly Met Arg Met Asp Leu Trp Lys Ser Arg Tyr
233          245         250         255
235 Lys Thr Phe Asp Glu Leu Tyr Leu Tyr Cys Tyr Tyr Val Ala Gly Thr
236          260         265         270

```

Input Set : A:\Mo6314 sequence-TEXT.txt  
Output Set: N:\CRF3\08092001\I847081A.raw

238 Val Gly Leu Met Ser Val Pro Val Met Gly Ile Ala Pro Glu Ser Lys  
239 275 280 285  
241 Ala Thr Thr Glu Ser Val Tyr Asn Ala Ala Leu Ala Leu Gly Leu Ala  
242 290 295 300  
244 Asn Gln Leu Thr Asn Ile Leu Arg Asp Val Gly Glu Asp Ala Arg Arg  
245 305 310 315 320  
247 Gly Arg Val Tyr Leu Pro Gln Asp Glu Leu Ala Gln Ala Gly Leu Ser  
248 325 330 335  
250 Asp Glu Asp Ile Phe Ala Gly Arg Val Thr Asp Lys Trp Arg Asn Phe  
251 340 345 350  
253 Met Lys Lys Gln Ile Gln Arg Ala Arg Lys Phe Phe Asp Glu Ser Glu  
254 355 360 365  
256 Lys Gly Val Thr Glu Leu Asp Ser Ala Ser Arg Trp Pro Val Leu Thr  
257 370 375 380  
259 Ala Leu Leu Leu Tyr Arg Lys Ile Leu Asp Glu Ile Glu Ala Asn Asp  
260 385 390 395 400  
262 Tyr Asn Asn Phe Thr Arg Arg Ala Tyr Val Ser Lys Pro Lys Lys Leu  
263 405 410 415  
265 Leu Thr Leu Pro Ile Ala Tyr Ala Lys Ser Leu Val Pro Pro Asn Arg  
266 420 425 430  
268 Thr Ser Ser Pro Leu Ala Lys Thr  
269 435 440

*Critical Error*  
*n at location 51 is not*  
*described.*

273 <210> SEQ ID NO: 3

274 <211> LENGTH: 1712

275 <212> TYPE: DNA

276 <213> ORGANISM: Nicotiana tabacum

278 <220> FEATURE:

279 <221> NAME/KEY: CDS

280 <222> LOCATION: (333)..(1565)

282 <220> FEATURE:

W--> 283 <221> NAME/KEY: Xaa — noncritical Error should be unknow or unsure  
284 <222> LOCATION: 135, 139 OK need to add location 51

285 <223> OTHER INFORMATION: Xaa is unknown or other OK

*Error*

W--> 286 <400> SEQUENCE: 3

W--> 287 cttgaagagt agcagcagca agcaagahaa ttcaaagtggg ctatttbkka naagccatttg 60  
288 ttacmagara attaagaagc caagamacag gttattttct acttgagtya ggaaaaagttg 120  
291 gttgcctta tttgtggct ttttataatc tttttccac aagggaaaagt gggtatattc 180  
293 ttgaaaagtgg attttagactc tagtggaat ctactaggag taaattttatt aatttttat 240  
295 aaattaagca gaggaaggaa ggaaaacagaa aacagaaaagt aagacaaaaa accttggaaat 300  
297 tgtttagaa agccaagggtt ttccgttca aa atg tct gtt gcc ttg tta tgg 353  
298 Met Ser Val Ala Leu Leu Trp  
299 1 5  
301 gtt gtt tca cct tgt gaa gtc tca aat ggg aca gga ttc ttg gat tca 401  
303 Val Val Ser Pro Cys Glu Val Ser Asn Gly Thr Gly Phe Leu Asp Ser  
304 10 15 20  
306 gtc cgg gag gga aac cgg gtt ttt gat tcg tcg agg cat agg aat tta 449  
308 Val Arg Glu Gly Asn Arg Val Phe Asp Ser Ser Arg His Arg Asn. Leu  
309 25 30 35  
311 gtg tgc aat gag aga aac aag aga ggt gtg aaa caa agg tgg aat ttt 497

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/847,081A

DATE: 08/09/2001

TIME: 12:44:02

Input Set : A:\Mo6314 sequence-TEXT.txt  
 Output Set: N:\CRF3\08092001\I847081A.raw

313	Val	Cys	Asn	Glu	Arg	Asn	Lys	Arg	Gly	Val	Lys	Gln	Arg	Trp	Asn	Phe	
314	40						45				50					55	
316	ggt	tct	gta	agg	tct	gtc	atg	gtg	gct	aca	ccg	gcg	gga	gaa	atg	gcg	545
318	Gly	Ser	Val	Arg	Ser	Ala	Met	Val	Ala	Thr	Pro	Ala	Gly	Glu	Met	Ala	
319							60			65					70		
321	acg	atg	aca	tca	gaa	cag	atg	gtt	tat	gat	gtg	gtt	tta	aaa	caa	gca	593
323	Thr	Met	Thr	Ser	Glu	Gln	Met	Val	Tyr	Asp	Val	Val	Leu	Lys	Gln	Ala	
324							75			80					85		
326	gct	tta	gtg	aag	agg	cag	ttg	aga	tct	gct	gat	gat	tta	gaa	gtg	aag	641
328	Ala	Leu	Val	Lys	Arg	Gln	Leu	Arg	Ser	Ala	Asp	Asp	Leu	Glu	Val	Lys	
329							90			95					100		
331	ccg	gag	atc	cct	ctc	ccc	ggg	aat	ttg	agc	ttg	ttg	agt	gaa	gca	tat	689
333	Pro	Glu	Ile	Pro	Leu	Pro	Gly	Asn	Leu	Ser	Leu	Leu	Ser	Glu	Ala	Tyr	
334							105			110					115		
336	gat	agg	tgt	agt	gaa	gta	tgt	gca	gag	tat	gca	aag	aca	ttt	tac	tth	737
W-->	338	Asp	Arg	Cys	Ser	Glu	Val	Cys	Ala	Glu	Tyr	Ala	Lys	Thr	Phe	Tyr	Xaa
339	120						125				130					135	
341	gga	acc	atg	yta	atg	act	cca	gag	aga	aga	agg	gct	att	tgg	gca	ata	785
W-->	343	Gly	Thr	Met	Xaa	Met	Thr	Pro	Glu	Arg	Arg	Arg	Ala	Ile	Trp	Ala	Ile
344							140			145					150		
346	tat	gtg	tgg	tgc	agg	aga	aca	gat	gaa	ctt	gtt	gat	ggc	cca	aac	gca	833
348	Tyr	Val	Trp	Cys	Arg	Arg	Thr	Asp	Glu	Leu	Val	Asp	Gly	Pro	Asn	Ala	
349							155			160					165		
351	tca	cat	att	aca	ccc	caa	gcc	tta	gat	agg	tgg	gaa	gac	cgg	ctt	gaa	881
353	Ser	His	Ile	Thr	Pro	Gln	Ala	Leu	Asp	Arg	Trp	Glu	Asp	Arg	Leu	Glu	
354							170			175					180		
356	gat	gtt	tcc	agc	ggg	cga	cca	ttt	gat	atg	ctc	gat	gct	gct	ttg	tcc	929
358	Asp	Val	Phe	Ser	Gly	Arg	Pro	Phe	Asp	Met	Leu	Asp	Ala	Ala	Leu	Ser	
359							185			190					195		
361	gat	act	gtt	tcc	aag	ttt	cca	gtt	gat	att	cag	ccg	tcc	aga	gat	atg	977
363	Asp	Thr	Val	Ser	Lys	Phe	Pro	Val	Asp	Ile	Gln	Pro	Phe	Arg	Asp	Met	
364							200			205					215		
366	att	gaa	gga	atg	cgt	atg	gac	t <sub>tg</sub>	agg	aag	tca	aga	tat	aga	aac	ttt	1025
368	Ile	Glu	Gly	Met	Arg	Met	Asp	Leu	Arg	Lys	Ser	Arg	Tyr	Arg	Asn	Phe	
369							220			225					230		
371	gat	gag	ctt	tac	ctc	tat	tgt	tat	tac	gtt	gct	ggt	acg	gtt	ggg	ttg	1073
373	Asp	Glu	Leu	Tyr	Leu	Tyr	Cys	Tyr	Tyr	Val	Ala	Gly	Thr	Val	Gly	Leu	
374							235			240					245		
376	atg	agt	gtt	cca	att	atg	ggt	att	gca	cct	gat	tca	aag	gca	aca	aca	1121
378	Met	Ser	Val	Pro	Ile	Met	Gly	Ile	Ala	Pro	Asp	Ser	Lys	Ala	Thr	Thr	
379							250			255					260		
381	gag	agc	gta	tat	aat	gca	gct	ttg	gct	tta	gga	atc	gca	aat	caa	cta	1169
383	Glu	Ser	Val	Tyr	Asn	Ala	Ala	Leu	Ala	Leu	Gly	Ile	Ala	Asn	Gln	Leu	
384							265			270					275		
386	acg	aac	ata	ctc	aga	gat	gtt	gga	gaa	gat	gcc	aga	aga	gga	aga	gtc	1217
388	Thr	Asn	Ile	Leu	Arg	Asp	Val	Gly	Glu	Asp	Ala	Arg	Arg	Gly	Arg	Val	
389							280			285					295		
391	tac	tta	cct	caa	gat	gaa	tta	gca	cag	gca	ggt	ctc	ttc	gac	gat	gac	1265
393	Tyr	Leu	Pro	Gln	Asp	Glu	Leu	Ala	Gln	Ala	Gly	Leu	Phe	Asp	Asp	Asp	

**VERIFICATION SUMMARY**  
PATENT APPLICATION: US/09/847,081A

DATE: 08/09/2001  
TIME: 12:44:04

Input Set : A:\Mo6314 sequence-TEXT.txt  
Output Set: N:\CRF3\08092001\I847081A.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:283 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3  
L:286 M:283 W: Missing Blank Line separator, <400> field identifier  
L:287 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:338 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:343 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:442 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:4  
L:472 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:892 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:5